

State-Dependent Learned Valuation Drives Choice in an Invertebrate

Lorena Pompilio,* Alex Kacelnik,† Spencer T. Behmer‡

Humans and other vertebrates occasionally show a preference for items remembered to be costly or experienced when the subject was in a poor condition (this is known as a sunk-costs fallacy or state-dependent valuation). Whether these mechanisms shared across vertebrates are the result of convergence toward an adaptive solution or evolutionary relicts reflecting common ancestral traits is unknown. Here we show that state-dependent valuation also occurs in an invertebrate, the desert locust *Schistocerca gregaria* (Orthoptera: Acrididae). Given the latter's phylogenetic and neurobiological distance from those groups in which the phenomenon was already known, we suggest that state-dependent valuation mechanisms are probably ecologically rational solutions to widespread problems of choice.

Animal decision-making is often modeled using the assumption that choices are based on the fitness consequences that each choice yields. Fitness gains, in turn, depend on both the intrinsic properties of the options and the state of the subject at the time of the choice. Recently, however, studies in humans and other vertebrates (1, 2) have shown that understanding the adaptive significance of learning mechanisms may be the key to progress in functional modeling of decision-making, because preferences more closely reflect the subject's state at the time of learning than at the time of choice. Classical learning models (3) do not address the subject's state, but recent treatments of evaluative incentive behavior do (4) and are compatible with the approach taken here.

A recent theoretical model linking learning to decision-making (5) proposes that anomalies of choice behavior in which past investments rather than expected returns dominate preference (examples include Sunk Costs, Work Ethics, and the Concorde Fallacy) result from a decelerated function of value (fitness or utility) versus objective payoff, combined with a mechanism of choice that is dependent on the remembered benefit previously yielded by each option (Fig. 1). Although some utility functions can be accelerated or sigmoid, because "desperados" in dire states would accrue less marginal gains from resources than would better-off individuals, most surviving organisms operate beyond this extreme zone, and hence the assumption of decelerated gains has very wide justification. In summary, if two sources (L and H) yielding the same objective payoff ($M_L = M_H$; M , magnitude) are systematically

encountered when the individual is in different states (low or high reserves for L and H, respectively), then the source encountered when needs are greater (L) will yield larger value gains ($V_L > V_H$). According to the model, although gains depend jointly on payoff magnitude and present state, it is the remembered gains, rather than remembered payoff magnitudes or states, that drive future preferences.

The adaptive advantages of such a mechanism are not obvious because, at least under experimental conditions, they can produce irrational preferences: Starlings can prefer a more delayed over a more immediate reward even when having explicit knowledge of the delays involved (6), and rats can frantically operate a lever or chain that causes food or water rewards even when being neither hungry nor thirsty (7). Supporting evidence for incentive or state-dependent learning comes from the mammal and bird species that have been studied so far, but there is an open question as to whether this mechanism of learned value assignment was an early vertebrate acquisition or a wider phenom-

enon perhaps universally present because it confers selective advantages.

We tested whether such state-dependent valuation learning occurs in a grasshopper, an animal with a simpler nervous system (8) than that of the vertebrates in which these effects are known. Grasshoppers make particularly good test subjects for studying and modeling individual decision-making because they forage for themselves and are capable of learning (9, 10). Additionally, much is known about how changes in their nutritional state affect their feeding behavior (11).

We manipulated nutritional state both at the time of learning and at the time of preference testing. We trained grasshoppers so that they encountered each of two options under different nutritional states: low (option L) and high (option H). Each option consisted of an odor (lemon grass or peppermint) paired during learning with a food item (a small piece of seedling wheat). Food items were of the same size and quality in both options, and each odor was always associated with the same state for each subject. Individuals received an equal number of reinforced trials with each option over a 3-day training regime (fig. S1). After training was completed, individual grasshoppers were presented with a choice between the two options. Half of the subjects had the test in the low state and the other half in the high state (12).

We considered four possible outcomes. The first of these, Magnitude Priority, states that if choices depend on the intrinsic properties of the options, no systematic preference will be observed between odors because the food items were identical. The second, Value Priority, states that if choices are controlled by past gains, preference should be for the option experienced in the low state during training,

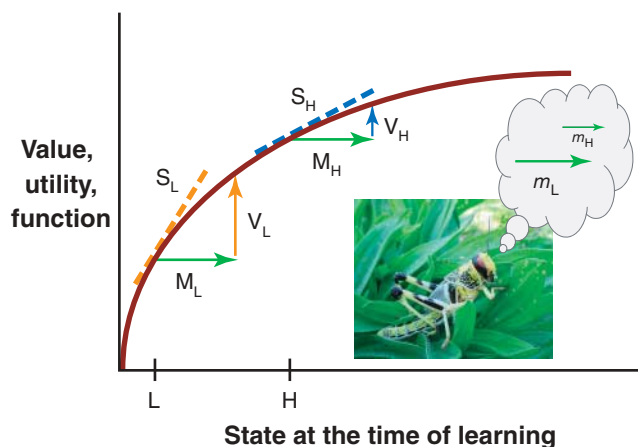


Fig. 1. Putative mechanisms of valuation learning as a function of a subject's state. The ordinate is a currency that is assumed to correlate to adaptive value, and the abscissa is a metric of objective state, here assumed to be the level of accumulated reserves. The plot illustrates consequences for a subject that encounters two food sources (L and H), each when the subject is in either of two states: low or high, respectively. The magnitudes of the outcomes are labeled M_L and M_H and are represented as arrows causing positive state displacements. The value (or benefit) of each outcome (V_L and V_H) is the vertical displacement that corresponds to each change in state. The first derivative (marginal rate) of the value-versus-state function at each initial state is indicated by the slopes of the tangents S_L and S_H .

The inset shows that the subject's representation of the magnitude of rewards (m) may differ from the objective metrics of the outcomes (in the example, $M_L = M_H$ but $m_L > m_H$). Models of learning may use M , S , V , or m as being directly responsible for value assignment.

Department of Zoology, University of Oxford, South Parks Road, Oxford OX1 3PS, UK.

*Present address: Departamento de Ecología, Genética y Evolución, Facultad de Ciencias Exactas y Naturales, Universidad de Buenos Aires, Pabellón II Ciudad Universitaria, C1428EHA, Buenos Aires, Argentina.

†To whom correspondence should be addressed. E-mail: alex.kacelnik@zoo.ox.ac.uk

‡Present address: Department of Entomology, Texas A&M University, College Station, TX 77843-2475, USA.

regardless of state at the time of choice. The third, State Priority, stipulates that if options are valued by association with the desirability of the state they evoke, the option preferred should be that experienced in the high state during training, regardless of state during choice. The fourth, State-Option Association, stipulates that choice should favor the source met under the same state at the time of training. Thus, subjects may choose option L under state low and option H under state high.

A majority of the grasshoppers preferred option L (the stimulus to which the grasshoppers were trained when in a state of low reserves) regardless of their state at the time of testing (Fig. 2). Averaged across all test subjects, the mean preference (\pm SE) for option L was 0.71 ± 0.06 . These results indicate a significant preference for option L ($t_{1,15} = 3.60$, $P < 0.01$; one sample t test against indifference). Preference was not affected by the state of the subject at the time of testing (Fig. 2) or by odor bias, and the state-by-odor interaction was not significant [analysis of variance (ANOVA): $F_{1,15} = 0.09$, $P > 0.77$; $F_{1,15} = 0.01$, $P > 0.92$; and $F_{1,15} = 0.01$, $P > 0.91$, respectively]. There also was no left arm–right arm positioning effect (paired-samples t test: $t_{1,15} = 0.17$, $P > 0.86$). Next, we considered whether the speed of learning during training, as measured by latencies to contact and eat the reward, might have anticipated the preference results. A repeated-measures ANOVA indicated that latencies to start eating decreased across the 3 days of training ($F_{2,15} = 15.00$, $P < 0.01$; fig. S2), but averaged over time, the latencies between the L and H options were similar ($F_{1,15} = 0.08$, $P > 0.78$), and no significant option-by-day interaction was observed

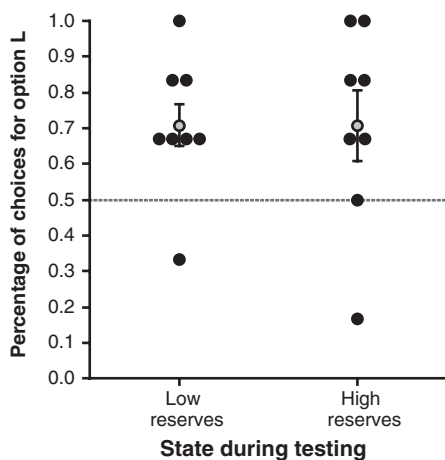


Fig. 2. Percentage of choices for option L exhibited by each subject under the two different states (low and high reserves) during testing. The figure shows individual data points (black circles) and means (\pm SE) (gray circles) with respect to indifference (dashed line). For both states, the percentage of choices for option L was significantly higher than indifference.

($F_{2,15} = 0.91$, $P > 0.42$). Finally, we considered the possibility that each grasshopper preferred the option for which it had a shorter latency during training (regardless of whether shorter latencies were exhibited for the L or H option) (12). We found, however, no association between latencies and choices (Pearson's correlation index, $r = 0.19$, $P = 0.48$).

Our experiment supports the idea that in this insect, the benefit gained at the time of training affects later preference even when the magnitudes of the rewards are equal (namely, Value Priority). The Value Priority outcome can be mediated by two very different mechanisms that could be labeled Perception Distortion and Remembered Value. Perception Distortion states that the energetic state at the time of training influences the distal mechanisms of perception, so that the memory of the properties of the options is altered: Equal payoffs are perceived as being different (In Fig. 1, $m_L > m_H$). Under the Remembered Value mechanism, the memory for the magnitudes is accurate, but the animal attaches different subjective attractiveness to each option, depending on its state while learning. These considerations may apply to similar anomalies of decision-making in all animals, including humans.

In grasshoppers, there is evidence favoring the Perception Distortion mechanism, because preference for the option experienced when reserves are low could be explained by peripheral gustatory responses underlying feeding behavior. In these animals, as time since the last feed increases, nutrient levels in the hemolymph drop, and as a consequence, mouthpart taste receptors become increasingly sensitive to key depleted nutrients (13, 14). This means that at a neurological level, a grasshopper with low reserves will receive greater feedback when it contacts a food item (15–17). Similarly, through digestive adaptations, individuals may extract more nutrients from identical food items when in greater need (18), and later choices may be governed by the memory of the postabsorptive gain (or the sensory adaptation consequent on the gains) and not of the objective features of the food items. The latter route for Perception Distortion could in theory also apply to vertebrates, but the available evidence does not point in this direction, at least for starlings. In their case, peripheral adjustments leading to either distorted representations or distorted perceptions due to rapid absorptive adaptations are both unlikely. This is because in learned valuation effects, starlings' preference between equally delayed rewards is not accompanied by alterations in the pecking rate (suggesting that neither the perception of the magnitude of the reward nor timing was altered) (1, 5, 6).

Thus, although similar behavioral outcomes are observed in starlings and grasshoppers, it is possible that different underlying mechanisms drive state-dependent learned valuation in each species. This difference supports the view that state-dependent learned valuation has intrinsic,

although not yet identified, adaptive advantages and has probably emerged and persisted in distant species via convergent evolution.

State-dependent valuation may be computationally more efficient than remembering the attributes of each option and weighing them against current nutritional state. This may reduce errors and help when decisions need to be made quickly and where neural constraints limit the amount of information that can be processed (19, 20). State-dependent valuation can cause suboptimal choices if there is a difference between the choice circumstances and the circumstances for learning about each option, as in our experiment. In particular, for there to be a cost, there must be a correlation between state and the probability of encounter with each option when options are met singly. This would occur because when these options are met simultaneously and a choice takes place, information about past gains could be misleading. Outside these probably rare circumstances, the mechanisms do not favor suboptimal alternatives (21). It could be argued that even if state-dependent valuation causes frequent and costly suboptimal choices in nature, it may persist because of neural or psychological constraints or because the cost associated with the development of a different mechanism is higher than the cost of using such a metric. The latter possibility cannot be discarded, but we do not favor it as a working hypothesis.

Ultimately, it would be ideal to measure the prevalence of different learning and choice circumstances in natural environments, but until that becomes possible, progress can be made by modeling the theoretical ecological worlds under which state-dependent valuation would be evolutionarily stable when used in competition with animals that form preferences based on the absolute properties of their options.

References and Notes

1. B. Marsh, C. Schuck-Paim, A. Kacelnik, *Behav. Ecol.* **15**, 396 (2004).
2. T. S. Clement, J. R. Feltus, D. H. Kaiser, T. R. Zentall, *Psychon. Bull. Rev.* **7**, 100 (2000).
3. R. A. Rescorla, A. R. Wagner, in *Classical Conditioning II: Current Research and Theory*, A. H. Black, W. F. Prokasy, Eds. (Appleton-Century-Crofts, New York, 1972), pp. 64–99.
4. B. W. Balleine, in *The Behavior of the Laboratory Rats: A Handbook with Tests*, I. Q. W. Kolb, Ed. (Oxford Univ. Press, Oxford, 2006), pp. 436–446.
5. A. Kacelnik, B. Marsh, *Anim. Behav.* **63**, 245 (2002).
6. L. Pompilio, A. Kacelnik, *Anim. Behav.* **70**, 571 (2005).
7. A. Dickinson, B. W. Balleine, in *Steven's Handbook of Experimental Psychology*, vol. 3, *Learning, Motivation and Emotion*, ed. 3, H. Pashler, R. Gallistel, Eds. (Wiley, New York, 2002), pp. 497–533.
8. M. Burrows, *The Neurobiology of an Insect Brain* (Oxford Univ. Press, Oxford, 1996).
9. R. Dukas, E. A. Bernays, *Proc. Natl. Acad. Sci. U.S.A.* **97**, 2637 (2000).
10. S. T. Behmer, C. E. Belt, M. S. Shapiro, *J. Exp. Biol.* **208**, 3463 (2005).
11. S. J. Simpson, D. Raubenheimer, *Adv. Study Behav.* **29**, 1 (2000).
12. Materials and methods are available as supporting material on Science Online.
13. S. J. Simpson, M. S. J. Simmonds, W. M. Blaney, S. James, *Appetite* **17**, 141 (1991).

14. S. J. Simpson, C. L. Simpson, *J. Exp. Biol.* **168**, 269 (1992).
 15. J. D. Abisgold, S. J. Simpson, *J. Exp. Biol.* **135**, 215 (1988).
 16. C. L. Simpson, S. Chyb, S. J. Simpson, *Entomol. Exp. Appl.* **56**, 259 (1990).
 17. R. F. Chapman, *Annu. Rev. Entomol.* **48**, 455 (2003).
 18. R. M. Sibly, in *Physiological Ecology: An Evolutionary Approach to Resource Use*, T. C. P. Calow, Ed. (Blackwell, Oxford, 1981), pp. 109–139.
 19. L. A. Real, *Am. Nat.* **140**, 5108 (1992).
 20. E. A. Bernays, *Annu. Rev. Entomol.* **46**, 703 (2001).
 21. G. Gigerenzer, *Behav. Brain Sci.* **27**, 336 (2004).
 22. We thank C. Nadell and C. Spatola Rossi for help during data collection; C. Schuck-Paim and M. S. Shapiro for helpful discussion; and T. Dickinson, P. Harvey, T. Reil, O. Lewis, and two anonymous referees for valuable feedback on the manuscript.

Supporting Online Material

www.sciencemag.org/cgi/content/full/311/5767/1613/DC1
 Materials and Methods
 Figs. S1 and S2
 References

15 December 2005; accepted 1 February 2006
 10.1126/science.1123924

An Equivalence Principle for the Incorporation of Favorable Mutations in Asexual Populations

Matthew Hegreness,^{1,2*} Noam Shores,^{1*} Daniel Hartl,² Roy Kishony^{1,3†}

Rapid evolution of asexual populations, such as that of cancer cells or of microorganisms developing drug resistance, can include the simultaneous spread of distinct beneficial mutations. We demonstrate that evolution in such cases is driven by the fitness effects and appearance times of only a small minority of favorable mutations. The complexity of the mutation-selection process is thereby greatly reduced, and much of the evolutionary dynamics can be encapsulated in two parameters—an effective selection coefficient and effective rate of beneficial mutations. We confirm this theoretical finding and estimate the effective parameters for evolving populations of fluorescently labeled *Escherichia coli*. The effective parameters constitute a simple description and provide a natural standard for comparing adaptation between species and across environments.

Spontaneous beneficial mutations are the fuel for adaptation, the source of evolutionary novelty, and one of the least understood aspects of biology. Although adaptation is everywhere—cancer invading tissues, bacteria escaping drugs, viruses switching from livestock to humans—beneficial mutations are notoriously difficult to study (1, 2). Theoretical and experimental advances have been made in recent years by focusing on the distribution of fitness effects of spontaneous beneficial mutations (3–8). Mapping the options for improvement available to single organisms, however, is insufficient for understanding the adaptive course of an entire population, especially in asexual populations of microorganisms or cancer cells where multiple mutations often spread simultaneously (9–16). Here, we use modeling and experimental results to show that the seeming additional complication of having multiple lineages competing within a population leads in fact to a drastic simplification: Regardless of the distribution of mutational effects available to individuals, a population's adaptive dynamics can be approximated by an equivalent model in which all favorable mutations confer the same fitness advantage, which we call the effective selection coefficient. We provide experimental estimates of the effective selection coefficient

and the corresponding effective rate of beneficial mutations for laboratory populations of *Escherichia coli*, and we demonstrate the predictive power of these effective parameters.

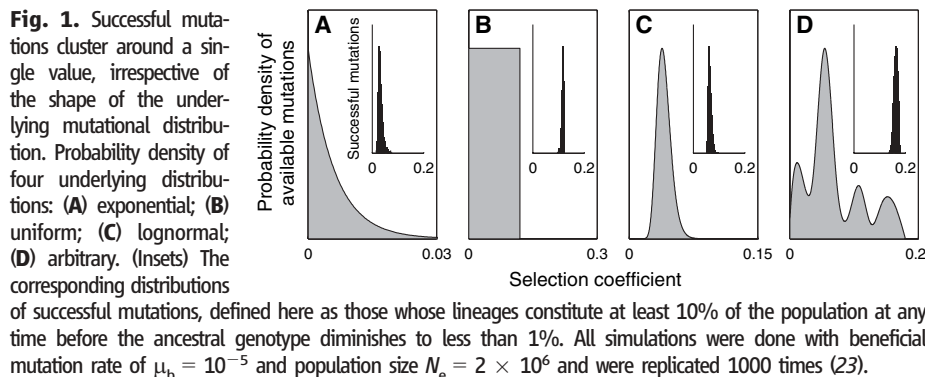
First, we use numerical simulations to demonstrate the simplification that emerges in a population large enough and a mutation rate high enough that clonal interference (17–19)—competition among lineages carrying favorable mutations—is common. In an evolving population, most beneficial mutations are rapidly lost to random genetic drift (20, 21). Of the remaining mutant lineages, some increase in frequency slightly, only to decline as more fit lineages appear and expand in the population (10, 16, 17, 22). The evolutionary path taken by the population as a whole is determined by successful mutations that escape stochastic loss and whose frequencies rise above some minimal level. Using a population genetics model that includes mutation, selection, drift, as well as clonal interference (23), we explore the distribution of these successful mutations for several underlying distributions of

beneficial mutations (Fig. 1), including an exponential distribution as suggested by Gillespie's (8) and Orr's (3) use of extreme value theory.

The salient feature of Fig. 1 is that very dissimilar underlying distributions—exponential, uniform, lognormal, even an arbitrary distribution—all yield a similar distribution of successful mutations (24). Moreover, the distribution of successful mutations has a simple form, peaked around a single value. This fitness value is typical of those mutations whose effects are not so small that they are lost through competition with more fit lineages, but are also not so large that they are impossibly rare. The unimodal shape motivates the hypothesis that an equivalent model that allows mutations with only a single selective value might approximate the behavior of the entire distribution of beneficial mutations.

We investigate whether the adaptive dynamics observed in evolving *E. coli* populations can be reproduced by an equivalent model with only a single value, a Dirac delta function of mutational effects. We rely on a classic strategy for characterizing beneficial mutations in coevolving subpopulations that differ initially only by selectively neutral marker. The spread of mutations is monitored through changes in the marker ratio (22, 25–29). Our experimental technique uses constitutively expressed variants of GFP (green fluorescent protein)—YFP (yellow fluorescent protein) and CFP (cyan fluorescent protein)—as neutral markers. All experimental populations start with equal numbers of YFP and CFP *E. coli* cells (N_Y and N_C) and evolve for 300 generations through serial transfers while adapting to glucose minimal medium.

The expected behavior of the marker-ratio trajectories depends upon the rate at which beneficial mutations appear in a population. When beneficial mutations are rare, mutant lineages arise and fix one at a time (8, 17). The spread of each individual mutant lineage shows as a line of



¹Bauer Center for Genomics Research, Harvard University, Cambridge, MA 02138, USA. ²Department of Organic and Evolutionary Biology, Harvard University, Cambridge, MA 02138, USA. ³Department of Systems Biology, Harvard Medical School, Boston, MA 02115, USA.

*These authors contributed equally to this work.

†To whom correspondence should be addressed. E-mail: roy_kishony@hms.harvard.edu